

Übernahme von Dissertationen und Habilitationen von der DBT in PICA über OAI

Bereitstellung der Daten in der DBT zur Übernahme

Die DBT bietet eine OAI Schnittstelle für das Metadaten-Harvesting-Protokoll der Open Archive Initiative, über welche gezielt Daten von verschiedenen nationalen und internationalen Service Providern eingesammelt werden können. Die DBT ist ein bei der Open Archive Initiative registrierter Data Provider mit der Basis URL

<http://www.db-thueringen.de/servlets/OAIDataProvider> .

Voraussetzung für das Einsammeln von Daten aus der DBT über diese registrierte URL ist der Eintrag eines Dokumentes in die Kategorie „für Harvesting bereitgestellt“ der Klassifikation „OAI“. Diese Klassifikation kann in der DBT nur von Mitgliedern der Gruppe „eval“, in der Regel Mitarbeitern der Bibliotheken, gefüllt werden. Diese Kategorie erhalten alle Dokumente, natürlich auch Dissertationen und Habilitationen, die die Bibliotheken evaluiert hat und den Bedingungen einer Langzeitarchivierung genügen. Um alle neu in die DBT eingepflegten Dissertationen abzusammeln, muß eine Abfrage speziell nach diesem Set erfolgen, eine zweite Abfrage dann nach Habilitationen.

[http://www.db-thueringen.de/servlets/OAIDataProvider?
verb=ListRecords&metadataPrefix=oai_dc&from=2008-04-01&set=type:6](http://www.db-thueringen.de/servlets/OAIDataProvider?verb=ListRecords&metadataPrefix=oai_dc&from=2008-04-01&set=type:6)

Diese Anfrage liefert die Records aller seit 1.4.2008 in die DBT eingepflegten Dissertationen. Eine Abfrage mit „set=type:9“ die Records der neuen Habilitationen.

Falls nicht alle an der DBT beteiligten Einrichtungen die automatische Übernahme der Dissertationen in den GBV wünschen, muß der abgeholte OAI Bestand selektiert werden. Die Informationen zur Herkunft einer Dissertation ist im Header „<spec>origin: Einrichtung</spec>“ enthalten.

Solange es sich nur um bestimmte Dokumenttypen wie Dissertationen oder Habilitationen handelt ist diese Praxis akzeptabel. Anders verhält es sich, wenn wir generell einen bestimmten Bestand unabhängig vom Dokumenttyp von der DBT in PICA übernehmen wollen. Folgende Vorgehensweise wäre dann denkbar:

Es ist möglich, weitere OAI Instanzen mit anderen Einschränkungen zu bilden, über welche für bestimmte Projekte ganz andere Dokumentensets über OAI angeboten werden. Für die Übernahme eines solchen Sets aus der DBT wurde eine zweite Instanz mit der Basis URL <http://www.db-thueringen.de/servlets/OAIGBVData> eingerichtet. Diese Instanz ist auf Dokumente eingeschränkt, die der Kategorie „GBV“ der Klassifikation „OAI“ angehören. Die Syntax der Abfrage ist die gleiche, nur dass eine andere Basis URL gewählt werden muss.

[http://www.db-thueringen.de/servlets/OAIGBVData?verb=getRecord&identifier=oai:www.db-thueringen.de:
7779&metadataPrefix=oai_dc](http://www.db-thueringen.de/servlets/OAIGBVData?verb=getRecord&identifier=oai:www.db-thueringen.de:7779&metadataPrefix=oai_dc)

liefert einen bestimmten Record, in diesem Fall das Dokument mit der ID 7779 im Dublin Core Format.

Sollte ein Dokument automatisch vom GBV übernommen werden, muß die entsprechende OAI Kategorie „GBV“ vergeben werden. Es wäre dann immer nur eine Abfrage nötig:

http://www.db-thueringen.de/servlets/OAIGBVData?verb=ListRecords&metadataPrefix=oai_dc&from=2008-04-01

liefert z.B. die Records aller seit 1.4.2008 eingepflegten Dokumente, die ausgewählt wurden, um automatisch in PICA übernommen zu werden. Voraussetzung für eine solche Praxis ist allerdings, dass man für jeden Dokumenttyp festlegt, welche Kategorien aus der DBT übernommen, welche Kategorien automatisch erzeugt werden und welche eine Nachbearbeitung durch die Bibliothek erfahren, bevor sie in PICA sichtbar werden.

OAI Record im Dublin Core Format

Beispiel für einen OAI Record aus der DBT:

```
<?xml version="1.0" encoding="UTF-8"?>
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2008-04-10T10:24:22Z</responseDate>
  <request verb="GetRecord" identifier="oai:DBT:10296"
metadataPrefix="oai_dc">http://www.db-
thueringen.de/servlets/OAIDataProvider</request>
  <GetRecord>
    <record>
      <header>
        <identifier>oai:www.db-thueringen.de:10296</identifier>
        <timestamp>2008-04-04</timestamp>
        <setSpec>origin:Physikalisch-Astronomische-Fakultät</setSpec>
        <setSpec>type:6</setSpec>
        <setSpec>format:1</setSpec>
      </header>
      <metadata>
        <oai_dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
          <dc:language>eng</dc:language>
          <dc:title xml:lang="eng">Time-Resolved X-Ray Diffraction Study on
Superconducting YBa2Cu3O7 Epitaxially Grown on SrTiO3</dc:title>
          <dc:creator>Lübcke, Andrea</dc:creator>
          <dc:format>Text</dc:format>
          <dc:type>Dissertation</dc:type>
          <dc:description>This thesis addresses the lattice response of a compound
consisting of a thin superconducting lm and a single crystalline substrate to an
```

ultrafast breaking of superconductivity. The lattice response is investigated by means of time-resolved X-ray diffraction. As X-ray source a laser produced plasma is used that emits sub-ps Ka bursts. This thesis is organized in the following way: After an introduction to important issues of the physics of high-temperature superconductors and SrTiO₃ as well as their lattice response, the experimental and theoretical methods used in this work are briefly introduced. In the following section, the experiment is discussed in more detail. Further, results from both, static and time-resolved experiments are presented. For a theoretical description of the experimental results several models are developed.</dc:description>

<dc:contributor>Förster, Eckhart</dc:contributor>

<dc:contributor>Seidel, Peter</dc:contributor>

<dc:contributor>Garcia, Martin E.</dc:contributor>

<dc:subject>Röntgendiffraktometrie, Hochtemperatursupraleiter</dc:subject>

<dc:date>2008-04-04</dc:date>

<dc:identifier>urn:nbn:de:gbv:27-20080404-112250-5</dc:identifier>

<dc:identifier>http://www.db-thueringen.de/servlets/DocumentServlet?id=10296</dc:identifier>

<dc:rights>free</dc:rights>

</oai_dc:dc>

</metadata>

</record>

</GetRecord>

</OAI-PMH>

Workflow zur Erfassung von Dissertationen/Habilitationen in der DBT mit Übernahme der Daten in PICA

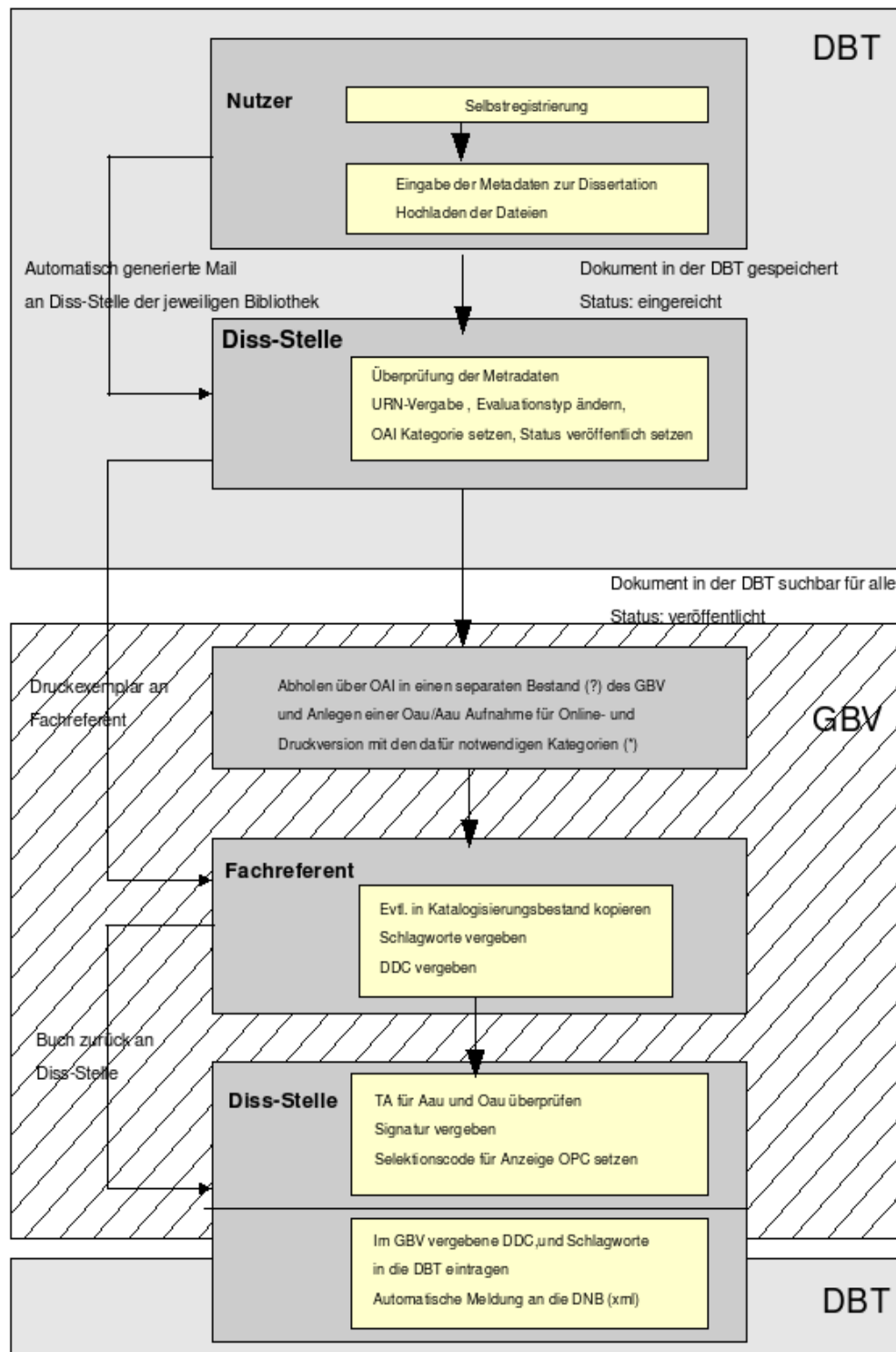


Abbildung 1: Workflow

Erklärung zum Workflow

Promoventen haben die Möglichkeit über die DBT Ihre Dissertationen/Habilitationen Online abzugeben, lediglich die geforderten 6 Pflichtexemplare und die Einverständniserklärung müssen in der Dissertationsstelle abgegeben werden.

Die folgende Erklärung beschränkt sich auf Dissertationen, die gleiche Prozedur gilt natürlich auch für Habilitationen.

1. Selbstregistrierung des Promoventen

Wenn der Promovent noch kein Autor der DBT ist, führt er zunächst eine Selbstregistrierung durch, gibt dafür seinen Vornamen, Nachnamen, Fachbereich, Telefonnummer und E-Mail Adresse sowie ein selbsternanntes Kennzeichen und ein Passwort ein. Der DBT Administrator erhält daraufhin eine E-Mail, dass sich ein Nutzer registriert hat, es besteht seinerseits kein Handlungsbedarf. Der so registrierte Nutzer ist ein Autor mit eingeschränkten Rechten, er darf nur Dokumente in der Kollektion Dissertationen/Habilitationen hochladen, seine Dokumente befinden sich immer im nicht veröffentlichten Zustand, was zur Folge hat, dass sie nicht suchbar sind, und die entsprechende Bibliothek wird per E-Mail über jede seiner Aktionen benachrichtigt. Sollte der Promovent schon Autor in der DBT sein, verwendet er dieses Kennzeichen für die Online Abgabe seiner Dissertation. Auch er lädt die Dissertation in die Kollektion Dissertationen/Habilitationen, in allen anderen Kollektionen, die er aufgrund seiner Rechte als Autor auch benutzen könnte ist der entsprechende Dokumenttyp „Dissertation“ und „Habilitation“ nicht verfügbar.

2. Dissertation Online abgeben

Der neue Nutzer kann sofort mit der Eingabe der Metadaten, sowie des Hochladens seiner Dissertation beginnen. Sobald die Metadaten als Dokument abgespeichert sind, sowie bei jeder weiteren Aktion, ändern, löschen, hochladen der Dateien etc. erhält die Dissertationsstelle eine entsprechende Mail über die Aktion und einen Link zum Dokument, über welchen sie wiederum die Metadaten überprüfen und ergänzen kann.

3. Überprüfung und Ergänzung der Metadaten durch die Dissertationsstelle

Die Dissertation/Habilitation befindet sich im Zustand „nicht veröffentlicht“. Die Dissertationsstelle überprüft die Metadaten, kann dazu mit dem Promoventen per E-Mail in Kontakt treten und gemeinsam Unstimmigkeiten beseitigen. In dieser Phase wird

- die URN vergeben und registriert,

- der Evaluationstyp der Dissertation auf Klasse A (Langzeitarchivierung) verändert und
- die Kategorie „für Harvesting bereitgestellt“ in der Klassifikation OAI gesetzt.

Nach endgültiger Fertigstellung wird das Dokument auf

- „veröffentlicht“ gesetzt, ist somit nach dem Abspeichern sofort suchbar .
- Außerdem wird das Datum „Geändert am:“ auf das aktuelle Datum geändert.
Mit dieser Aktion wird die Dissertation zum Abholen durch den Service Provider GBV bereitgestellt.
Mit diesen beiden Aktionen wird die Dissertation zum Einsammeln per OAI bereitgestellt.
- Das Druckexemplar wird an den Fachreferenten weitergereicht.

4. **Abholen über OAI**

Abhängig vom zeitlichen Rhythmus, der mit dem GBV vereinbart wird, wird das Dokument beim nächsten Abholen mit erfaßt und entweder in einem separaten Bestand des GBV oder direkt im Katalogisierungsbestand eine Aufnahme (Oax oder Oau?) angelegt. Auf diese Weise werden keine Dubletten im Katalogisierungsbestand erzeugt. Dabei werden die meisten Kategorien automatisch mit Attributen aus den OAI Records der DBT belegt, bzw. gemäß den Forderungen für Online Dissertationen eingefügt. Die Liste der bisherigen automatisch belegten Kategorien ist in Abbildung 2 im mittleren Feld beschrieben.

Beim derzeitigen Stand des Projektes fehlen teilweise noch Kategorien. Ziel ist eine möglichst vollständige Aufnahme mit Ausnahme der Kategorien 5010/5051 (DDC) sowie 5550 (Schlagworte). Der Exemplarsatz in Kategorie 7001 sollte mit dem Selektionscode „a“ angelegt werden. Überlegenswert wäre auch das gleichzeitige automatische Anlegen einer Titelaufnahme für das Druckexemplar (Aax oder Aau ?). Auch dort müßten nur wenige Kategorien händisch ergänzt werden.

5. **Bearbeitung der Oa? Aufnahme durch den Fachreferenten**

Der Fachreferent hat das Druckexemplar erhalten, somit ist ihm bekannt, dass in nächster Zeit eine neue Titelaufnahme bereitstehen wird. In Abhängigkeit von der Verfahrensweise der Übernahme in den Katalogisierungsbestand (extra Bestand oder nicht) muss er die Aufnahme in den Katalogisierungsbestand kopieren oder nicht. Er kann jetzt die Aufnahme vervollständigen, Schlagworte und DDC vergeben und anschließend das Druckexemplar an die Dissertationsstelle zurückreichen.

6. **Endgültige Bearbeitung der Aufnahme durch die Dissertationsstelle**

Durch die Zurückgabe des Druckexemplars wird der Dissertationsstelle signalisiert, dass nun die

Aufnahme endgültig fertiggestellt werden kann.

- Die Titelaufnahme für Aau und Oau wird überprüft
- eine Signatur vergeben und
- der Selectionscode für die Anzeige im OPC gesetzt.

7. Abschließende Änderung der Metadaten in der DBT

In der DBT können nun der im GBV vergebene DDC und die Schlagworte eingetragen werden. Die sachliche Erschließung ist noch abschließend zu klären.

Zum Schluss wird über die DBT per Mausklick die automatische Mail an die DNB generiert und verschickt.

Metadatenkonvertierung DBT - PICA

Metadaten DBT Einträge im spez. OAI PICA Katalogeintrag (2) Kategorienbeschreibung

Bestand in PICA (1)

Datum * dc:date	→	0500 Oax	0500 Oax	Bibliogr. Gattung
	→	1100 Jahr	1100 Jahr	Erscheinungsjahr
			1108 Elektronische Resource	Allgemeine Materialbenennung
			1110 ho	Code für Dok.typ(t)
			1110 o3	Code für Dok.typ
Sprache* dc:language	→	1500/1ger	1500/1ge	Sprachcode
			1700/1XA-DE	Erscheinungsland)
Autor* dc:creator	→	3000 Autor	2250 urn:nbn:de:gbv:27-....	URN
Gutachter* dc:contributor	→	3011 Gutachter	3000 Autor	
Gutachter* dc:contributor	→	3012 Gutachter		
Titel * dc:title	→	4000 Titel	4000 Titel	
Provenienz* im spec			4031 Jena: Thüringer ...	Ort, Datenbankanbieter
			4060 Größe in MB, Text	Umfangsang., spez.Material
			4061 graph. Darst	Illustr. Ausgabe
URN* dc:identifier	→	4083	4083 Verlinkung zum Volltext (3)	
URL* dc:identifier	→	4083	4083 Verlink. Frontpage DBT (3)	
Dokumenttyp* dc:type	→	4204@Jena, <docyp>< jahr>	4204 @Jena, Univ. <doctype>, < jahr>	HS-Schriftenvermerk
Beschreibung* dc:description	→	4207 Beschreibung	4207 Beschreibung	Abstract (kurz)
			4243 Druckausg., Titel mit Verlinkung	Horizontale Verknüpfung
DDG	→	nach Übernahme in PICA	5010 DDG oder 5051 (von der DNB)	
			5301	BKL
Stichworte dc:subject	→	nach Übernahme in PICA	5550 lokale Schlagworte	Einzel Schlagwort
			7001	Exemplarzählung
			7901	Bearbeitungsdatum

1 Über OAI aus der DBT gezogene Daten in einem Spezialbestand des GBV

2 Gewünschtes Ziel: aus 1 automatisch ergänzt zu fast vollständigem Titel mit Ausnahme der Kateg. 4060, 5010, 5550, diese werden durch die jeweilige Bibliothek ausgefüllt

3 Lizenzindikator '1' für frei setzen

Folgende Fragen müssen noch geklärt werden:

1. Der Lizenzindikator sollte in der Regel auf "1"=lizenzfrei gestellt werden
2. Der "dok-typ" im HSS-Vermerk kann aus dem Eintrag in der digitalen Bibliothek abgeleitet werden; Diss. / Habil.-Schr. / ...usw. Das wäre auch für spätere andere automatische Übernahmen (Monographien ...) wichtig.
3. Kann in der Kategorie 7901 (Bearbeitungsdatum) die Angabe einer vzg-Kennung (z.B. oai_ueb_vzg/1999)
4. Die Angabe der URL würden wir auch nur in der bibliographischen Ebene vorsehen und nicht zusätzlich als Wiederholung im Exemplarsatz, da der Zugang allgemeingültig ist.
5. Könnte mit der Eintragung des elektronischen Exemplars auch gleichzeitig die Titelaufnahme für die Druckschrift per OAI mit den dafür erforderlichen Kategorien angelegt werden?
6. Bezüglich der zu erfassenden Kategorien für Oa- und Aa-Aufnahme ist eine Anfrage an die VZG (Fr.Dr. Block/Frau Sylvia Langer) gestellt worden, um eine verbundeneinheitliche Katalogisierung zu gewährleisten.
7. Beim Sichten von Titeln zu derzeit noch per Hand eingegebenen Dissertationen sind wir darauf gestoßen, dass neben der eigenen Aufnahme ab einem bestimmten Zeitpunkt noch die Oax-Aufnahme der DNB steht. Beispiele für Suchanfragen im CBS:
f mak Oa? and taf "rolle der von willebra?"
f mak Oa? and taf "möglichkeiten und grenzen einer comp?"
f mak Oa? and taf "dünnschichtsysteme für die effektive?"
f mak Oa? and taf "hirnmorphologische veränderungen?"
 - Wie sind die organisatorischen Abläufe bei Dissertationen?
Gäbe es auch die Möglichkeit, vorhandene Titelaufnahmen, die in einem ordentlichen definierten Schema vorliegen, um die Informationen der DNB zu ergänzen?
 - Die Vergabe des Lizenzindikators in den DNB-URLs verstehen wir auch nicht so ganz, - warum "0" fuer lizenzpflchtig?
In unserer Digitalen Bibliothek sind die Dokumente frei zugänglich.
 - An den Oax-Aufnahmen sind teilweise auch Exemplarsätze anderer Bibliotheken.
(Was bedeutet eigentl. "add_edisex/1999"?)

Ulrike Krönert
Heidrun Möckel

23.4.2008